HHS Public Access

Author manuscript

Mol Genet Genomics. Author manuscript; available in PMC 2018 February 20.

Published in final edited form as:

Mol Genet Genomics. 2018 February ; 293(1): 225–235. doi:10.1007/s00438-017-1381-6.

Improved detection of genetic loci in estimated glomerular filtration rate and type 2 diabetes using a pleiotropic cFDR method

Hui-Min Liu[1], **Jing-Yang He**[1], **Qiang Zhang**[1], **Wan-Qiang Lv**[1], **Xin Xia**[1], **Chang-Qing Sun**[1], **Wei-Dong Zhang**[1], and **Hong-Wen Deng**[1,2]

[1]College of Public Health Zhengzhou University, No.100 Kexue Road, High-Tech Development Zone of States, Zhengzhou, People's Republic of China

[2]Department of Biostatistics and Data Science, Tulane Center of Bioinformatics and Genomics, Tulane University School of Public Health and Tropical Medicine, New Orleans, LA 70112, USA

Abstract

Genome-wide association studies (GWAS) have been shown to have the potential of explaining more of the "missing heritability" of complex human phenotypes by improving statistical approaches. Here, we applied a genetic-pleiotropy-informed conditional false discovery rate (cFDR) to capture additional polygenic effects associated with estimated glomerular filtration rate (creatinine) (eGFRcrea) and type 2 diabetes (T2D). The cFDR analysis improves the identification of pleiotropic variants by incorporating potentially shared genetic mechanisms between two related traits. The Q–Q and fold-enrichment plots were used to assess the enrichment of SNPs associated with eGFRcrea or T2D, and Manhattan plots were used for showing chromosomal locations of the significant loci detected. By applying the cFDR method, we newly identified 74 loci for eGFRcrea and 3 loci for T2D with the cFDR criterion of 0.05 compared with previous related GWAS studies. Four shared SNPs were detected to be associated with both eGFRcrea and T2D at the significant conjunction cFDR level of 0.05, and these shared SNPs had not been reported in previous studies. In addition, we used DAVID analysis to perform functional analysis of the shared SNPs' annotated genes and found their potential hidden associations with eGFRcrea and T2D. In this study, the cFDR method shows the feasibility to detect more genetic variants underlying the heritability of eGFRcrea and T2D, and the overlapping SNPs identified could be regarded as candidate loci that provide a thread of genetic mechanisms between eGFRcrea and T2D in future research.

Correspondence to: Wei-Dong Zhang; Hong-Wen Deng.

Electronic supplementary material The online version of this article (doi:10.1007/s00438-017-1381-6) contains supplementary material, which is available to authorized users.

Author contributions H-ML as the first author performed data analysis and wrote/revised the manuscript. JYH, QZ, XX, W-QL and C-QS provided advice and suggestions when we faced some problems during the data analysis process and revised the manuscript. W-DZ and H-WD are the co-corresponding authors. H-WD conceived and initiated this project, provided advice on experimental design, oversaw the implementation of the statistical method and finalized the manuscript revision. W-DZ revised the manuscript and oversaw the point to point response to reviewers in our rebuttal later.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Keywords

eGFRcrea; T2D; Pleiotropic; cFDR

Introduction

Failing to explain a substantial proportion of the heritability of complex phenotypes is often described as the “missing heritability” problem in traditional genome-wide association studies (GWAS) (Pei et al. 2014). Heritable complex diseases and traits, including estimated glomerular filtration rate (eGFR) and T2D, are widely believed to have underlying missing heritability in previous respective GWASs. Estimated glomerular filtration rate by serum creatinine (eGFRcrea) has been one of the crucial clinical techniques to estimate the kidney function decline (Okada et al. 2012). In standard GWAS analysis, the heritability of GFR has been estimated as 36–75% (O’Seaghdha and Fox 2012). However, the identified 16 SNPs by GWAS, e.g., by Köttgen et al., only accounted for 1.4% of the variability of eGFR (Regele et al. 2015). T2D is another chronic metabolic disorder that is characterized by high serum glucose, insulin resistance, and relative lack of insulin. Previous studies estimated that T2D affected over 387 million people with a worldwide prevalence of 8.3% in 2014 (Golay and Ybarra 2005; Karaderi et al. 2015). Heritability of T2D has been estimated ranging from 26 to 69%, but genes uncovered can only explain a small fraction of the heritability (Poulsen et al. 1999; Almgren et al. 2011).

Much effort has been made to excavate the “missing heritability” mainly by expanding sample sizes. However, the approach, although straightforward, is highly expensive (Stahl et al. 2012). Recently, cost-effective analytical methods, including summary statistics-based multivariate meta-analysis of GWAS using canonical correlation analysis (metaCCA) and Genetic analysis incorporating Pleiotropy and Annotation (GPA), have been developed to efficiently utilizing the existing datasets for improving the gene discovery (Chung et al. 2014; Cichonska et al. 2015, 2016). In the current study, a recently developed pleiotropic-informed approach conditional false discovery rate (cFDR) was widely adopted to improve the gene discovery in schizophrenia and bipolar disorder (Andreassen et al. 2013c), schizophrenia and cardiovascular-disease risk factors (Andreassen et al. 2013a), and successfully discovered novel loci which indicated that the pleiotropic loci exist between two related traits/diseases. Recently, we implemented the cFDR analyses and successfully identified pleiotropic variants in our groups (Zeng et al. 2016; Greenbaum et al. 2017; Lv et al. 2017; Peng et al. 2017; Zhou et al. 2017). Pleiotropy is defined as a single locus or gene affects more than one trait. It was conservatively estimated that 16.9% genes and 4.6% SNPs in the human genome have pleiotropic effects (Sivakumaran et al. 2011). Incorporating these pleiotropic effects between related diseases or traits may effectively enhance the ever-larger sample sizes across existing GWAS datasets and thus will be helpful in adding to the novel discoveries of genetic associations underlying the missing heritability (Andreassen et al. 2013a, c, 2014; Zeng et al. 2016; Greenbaum et al. 2017).

T2D and eGFRcrea are widely believed to have some common underlying polygenic architecture. Epidemiological studies suggested that eGFR changes during the development

of diabetes (Lorenzo et al. 2009), which links eGFR to insulin resistance (Chen et al. 2004; Lorenzo et al. 2009) and other potential abnormalities: increased renal gluconeogenesis (Eid et al. 2006) and activation of the renin–angiotensin system (Yamamoto et al. 2007). Genes SLC9B2 and VEGFA have been reported as genetic risk loci for eGFRcrea (Eremina et al. 2007; Deisl et al. 2013) or T2D (Sharma et al. 2011; Deisl et al. 2013) in different studies, respectively. Given the close relationship of two traits in epidemiological studies (Chen et al. 2004; Eid et al. 2006; Yamamoto et al. 2007; Lorenzo et al. 2009) and given a few common genes reported in respective GWAS reflect that eGFRcrea and T2D may share some potential pleiotropic genetic determination, further exploration by cFDR is warranted as to be done here.

In this study, we performed genetic-pleiotropy-informed cFDR to capture additional polygenic effects associated with eGFRcrea and T2D. This method utilizes summary statistics from two independent large GWAS meta-analysis datasets of eGFRcrea (Pattaro et al. 2016) and T2D (Mahajan et al. 2014), which will increase the power of gene discovery and improve gene detection of these two related traits for those loci with potential pleiotropic variants by effectively increasing samples sizes (Andreassen et al. 2013a, c, 2014).

Materials and methods

GWAS datasets

The GWAS summary statistics including *p* values of SNPs were obtained from two available datasets. The dataset for eGFRcrea was taken from a GWAS meta-analysis of 133,413 European-ancestry subjects from 48 individual studies, which was performed by CKDGen consortium published in 2015 (https://fox.nhlbi.nih.gov/CKDGen/) (Pattaro et al. 2016). To our knowledge, it is the largest reported eGFRcrea-associated GWAS study to date. The dataset for T2D was taken from a total of 26,488 cases and 83,964 controls from the trans-ethnic T2D GWAS meta-analysis published in 2014 (http://www.diagram-consortium.org/downloads.html) (Mahajan et al. 2014). Detailed information about the original data preparation and methods is provided in the following individual studies: the DIAbetes Genetics Replication and Meta-analysis (DIAGRAM) Consortium5 (European ancestry; 12,171 cases and 56,862 controls) (Morris et al. 2012), the Asian Genetic Epidemiology Network T2D (AGEN-T2D) Consortium11 (East Asian ancestry; 6952 cases and 11,865 controls) (Cho et al. 2012), the South Asian T2D (SAT2D) Consortium13 (South Asian ancestry; 5561 cases and 14,458 controls) (Kooner et al. 2011), and the Mexican American T2D (MAT2D) Consortium15 (Mexican and Mexican American ancestry; 1804 cases and 779 controls) (Parra et al. 2011). After applying appropriate quality filters, about ~ 2.5 million SNPs have been imputed in each datasets. Genomic control (GC) was conducted twice, one before and the other after the meta-analysis, the first was before preparation within individual studies for the meta-analysis and the second was after the meta-analysis for further limiting the possibility of false positives due to the pooling of data from various studies for meta-analysis (Mahajan et al. 2014; Pattaro et al. 2016). The detailed inclusion criteria and phenotype characteristics in different GWAS studies were described in the original publications (Mahajan et al. 2014; Pattaro et al. 2016).

Data preparation

Several steps were involved in the data preparation. First, we checked the European ancestry cohorts for overlapping individuals included in the GWAS samples. An upper bound for the amount of sample overlap is obtainable from the original publications by comparing the sub-study definitions and sample sizes for eGFRcrea and T2D. The details were illustrated in Supplementary Table S1 and we assume this will not affect the detecting power of the cFDR analysis (LeBlanc et al. 2016). Second, the SNPs were selected and clustered into independent loci through linkage disequilibrium (LD) (Andreassen et al. 2013b, 2014) pruning based on $r^2 > 0.2$ within a window of 50 SNPs. For each window, LD was calculated between each pair of SNPs and the SNP with smaller MAF between pairs of variants was removed. Then, the window slid 5 SNPs forward and the above pruning process was repeated until there were no pairs of SNPs with high LD. The dataset was pruned using the HapMap 3 genotypes CEU (http://www.sanger.ac.uk/resources/downloads/human/hapmap3.html) and there were 100,952 variants left for our analysis. Since GC has been applied in both individual GWAS studies and original meta-analysis studies (Mahajan et al. 2014; Pattaro et al. 2016), there is no need to reapply GC in our analysis.

Statistical analysis

Conditional quantile–quantile (Q–Q) and fold-enrichment plots for pleiotropic enrichment—Conditional Q–Q plots were constructed to evaluate the pleiotropic enrichment by conditioning the principal trait on the SNPs with varying strengths of association in the conditional trait (Andreassen et al. 2013c). The Q–Q plots show the observed distribution of p values plotted against the expected distribution of p values under the null hypothesis. In this study, Q–Q plots were presented by $-\log_{10}$ nominal p values plotted on the y-axis against $-\log_{10}$ empirical conditional p values [empirical cumulative distribution functions (cdfs)] plots on the x-axis based on varying levels of $p < 1$, $p < 0.1$, $p < 0.01$, $p < 0.001$ and $p < 0.0001$. Leftward deflections of the observed distribution from the projected null line reflect increased tail probabilities in the distribution of test statistics (Z scores) and, consequently, an overabundance ("enrichment") of low p values compared to that expected by chance. An increased deflection from the null line in the Q–Q plots indicates pleiotropic enrichment shared between the principal and conditional traits (Schork et al. 2013). Additionally, we conducted the fold-enrichment plots. The nominal p values [$-\log_{10}(p)$] are plotted on the x-axis, and fold enrichment in eGFRcrea as a function of T2D is plotted on the y-axis. We present fold-enrichment plots of nominal $-\log_{10}(p)$ values for eGFRcrea SNPs below the standard GWAS threshold of $p < 5 \times 10 - 8$ and for subsets of SNPs determined by the significance of their association with T2D and vice versa. As a supplement to confirm the pleiotropic enrichment effect, which can be completed by comparing the proportion of SNPs reaching each cutoff level in different significant groups with the group including all SNPs ($p = 1$). Pleiotropic enrichment is assessed by the degree of upward shift from the expected null line.

Conditional false discovery rate (cFDR) and conditional Manhattan plots—The cFDR method is an extension of standard FDR and was generated from the empirical Bayes method (Efron 2007). The cFDR method combines summary statistics to obtain the

probability of the association of the principal phenotype conditioned on the strength of association with conditional phenotype. cFDR was expressed as:

$$\mathrm{cFDR}(p_i|p_j)=\Pr(H_0^{(i)}|P_i \le p_i, P_j \le p_j)$$

The cFDR($p_i|p_j$) for each SNP was where eGFRcrea is the principal phenotype conditioned on the strength of association with T2D (eGFRcrea|T2D) and vice versa. Where p_i represents the strength of association of a random SNP with the 'principal phenotype', p_j is the strength of association of the same SNP with the 'conditional phenotype'. $H_0^{(i)}$ represents the null hypothesis that a given SNP is not associated with the principal trait. In the present study, cFDR was computed for each SNP where eGFRcrea is the principal phenotype conditioned on the strength of association withT2D (eGFRcrea|T2D) and vice versa (T2D|eGFRcrea).

The conditional Manhattan plots were adopted to display the chromosomal locations of all the loci associated with eGFRcrea based on T2D and vice versa. A criterion of $-\log_{10}(\mathrm{FDR}) > 1.3$ (corresponding to cFDR < 0.05) was used to select loci associated with eGFRcrea or T2D.

Conjunction cFDR and conjunction Manhattan plots—To detect pleiotropic loci shared between eGFRcrea and T2D, we computed the conjunction cFDR (ccFDR). It is defined as the probability that a given SNP has a false-positive association with both eGFRcrea and T2D. The ccFDR was computed as the maximum cFDR values of the two traits. SNPs with ccFDR < 0.05 will be considered significantly associated with both traits. Conjunction Manhattan plot was presented to illustrate the chromosomal locations of pleiotropic loci associated with both eGFRcrea and T2D. SNPs with $-\log_{10}(\mathrm{cFDR}) > 1.3$ (ccFDR < 0.05) are significantly associated with both eGFRcrea and T2D.

Functional annotation analysis for pleiotropic loci—Functional annotation for pleiotropic loci associated with eGFRcrea and T2D was conducted by DAVID Bioinformatics Resources 6.8 database, which was downloaded from the DAVID Consortium (https://david.ncifcrf.gov/home.jsp) (Huang et al. a, 2009b).

DAVID covers multiple data resources based on functional annotation analysis, and we employed GO, INTERPRO, KEGG_PATHWAY, OMIM_DISEASE and UP_KEYWORDS in this study. Using the functional annotation analysis, we characterized the trait-associated loci based on their known biological processes and molecular functions to understand the potential biological mechanisms behind the large list of discovered genes. This analysis allows us to validate our findings by determining functional annotation of the gene sets that are significantly associated with both eGFR and T2D. In the current study, gene sets with ccFDR < 0.05 for both eGFRcrea and T2D were performed in DAVID.

Results

Assessment of pleiotropic enrichment

Conditional Q–Q plots were presented to assess the pleiotropic enrichment between eGFRcrea and T2D in Fig. 1. Specifically, conditional Q–Q plots for eGFRcrea conditioned on T2D (eGFRcrea|T2D) and vice versa (T2D|eGFRcrea) are shown in Fig. 1a, b. As reflected in Fig. 1a, the leftward shift from the null line indicates association between SNPs with eGFRcrea conditioned on T2D for given p values. In Fig. 1b, (T2D|eGFRcrea), we obtained similar result. What's more, both in Fig. 1a, b, while we descended the thresholds at the level of $-\log_{10}(p) > 0$, $-\log_{10}(p) > 1$, $-\log_{10}(p) > 2$, $-\log_{10}(p) > 3$ and $-\log_{10}(p) > 4$ ($p < 1$, $p < 0.1$, $p < 0.01$, $p < 0.001$, $p < 0.0001$), there would be greater amounts of separation in the curves. The leftward shifts suggested strong level of enrichment and great proportion of associations for eGFRcrea and T2D. Pleiotropic enrichment effects can also be assessed by the degree of upward shift from the expected null line ($p = 1$) in the fold-enrichment plots, and the more upward shift means more fold enrichment. Based on the fold-enrichment plots in Fig. 1c, d, we observed approximately 15-fold increase for eGFRcrea and 19-fold increase for T2D in the proportion of SNPs reaching the genome-wide significance level of $-\log_{10}(p)$ − values > 7.3 ($p < 5 \times 10^{-8}$) when comparing the subset with the most stringent conditional association ($p = 0.001$) to the group with all SNPs ($p = 1$). In addition, the fold enrichment further reflects the association between eGFRcrea and T2D.

eGFRcrea loci identified with cFDR

We totally detected 137 SNPs significantly (cFDR < 0.05) associated with eGFRcrea conditioned on T2D, which were located on 22 different chromosomes (chr1–13, 15–21), and the results were presented by the conditional Manhattan plot in Fig. 2. Of these 137 SNPs, 71 SNPs had p values smaller than 1×10^{-5} (~ 51.8%), while 35 SNPs reached genome-wide significance at 5×10^{-8} (~ 25.5%) in the original meta-analysis study for eGFRcrea (Pattaro et al. 2016). We validated 10 SNPs which were reported to be associated with renal function (details displayed in the Supplementary Table S2) (Kottgen et al. 2010; Tin et al. 2013; Pattaro et al. 2016). Eight of these 10 SNPs were associated with GFR (Tin et al. 2013; Pattaro et al. 2016), and two were associated with chronic kidney disease (Kottgen et al. 2010; Pattaro et al. 2016). Another 53 SNPs had not been discovered to be eGFRcrea-associated loci, but their annotated genes have been verified to have a relationship with GFR (details displayed in the Supplementary Table S3) (Mahajan et al. 2016; Pattaro et al. 2016). The rest 74 SNPs and their annotated genes were newly found to be associated with eGFRcrea (details displayed in the Supplementary Table S4).

T2D loci identified with cFDR

Conditional Manhattan plot for T2D (cFDR < 0.05) showed a total of 19 SNPs significantly associated with T2D given their association with eGFRcrea, which were located on 11 chromosomes (chr1, 3–8, 10–11, 16–17) (Fig. 3). Of these 19 SNPs, 17 had p values smaller than 1×10^{-5}, while 6 SNPs reached genome-wide significance at 5×10^{-8} in the original meta-analysis for T2D (Mahajan et al. 2014). Of these significant SNPs, three have been validated significantly associated with T2D (details in the Supplementary Table S5) (Li et al.

2013; Mahajan et al. 2014). Another subset of 13 SNPs had not been reported to be associated with T2D but their annotated genes were mentioned as T2D associated in previous studies (details in the Supplementary Table S6) (Sim et al. 2011; Voight et al. 2011; Mahajan et al. 2014; Cook and Morris 2016). SNPs rs2881654 (*PPARG,* $p = 1.7 \times 10^{-7}$, cFDR = 0.0023), rs7079711 (*TCF7L2,* $p = 2.5 \times 10^{-7}$, cFDR = 0.0026) and rs11979110 (*KLF14,* $p = 1 \times 10^{-7}$, cFDR = 0.001) had been detected to be associated with T2D by cFDR, but might be missing under the strict GWAS standard of 5×10^{-8} (Mahajan et al. 2014). Above all, we identified three T2D-associated novel loci rs4699049 (*SLC9B2*), rs3843467 (*C5orf67*) and rs227375 (*MANBA*) and detailed information will be displayed in Supplementary Table S7.

Pleiotropic loci for both eGFRcrea and T2D

To identify the pleiotropic loci between eGFRcrea and T2D, we performed a conjunction cFDR (ccFDR) analysis (ccFDR < 0.05) and constructed a conjunction Manhattan plot for their chromosomal locations. In total, we detected four pleiotropic variants rs4699049 (*SLC9B2*), rs9381257 (*VEGFA* and *LINC01512*), rs227375 (*MANBA*) and rs1876602 (*MTNR1B* and *SLC36A4*) significantly associated with both eGFRcrea and T2D, which were located on three chromosomes (chr. 4, 6, 11) (Fig. 4; Table 1). These four SNPs are all found to be associated with both eGFRcrea and T2D. The annotated gene *VEGFA* by locus rs9381257 has been identified to be associated with eGFRcrea (rs9472135 in *VEGFA*) (Pattaro et al. 2016) or T2D (rs9472138 in *VEGFA*) (Mahajan et al. 2014), but no studies have reported that *VEGFA* was associated with both eGFRcrea and T2D earlier.

Functional annotation analysis for pleiotropic loci

A series of bioinformatics analyses in DAVID was conducted to explain the potential functions for the six genes (*MANBA, MTNR1B, SLC36A4, SLC9B2, LINC01512* and *VEGFA*) located in four pleiotropic loci. Five genes (*MANBA, MTNR1B, SLC36A4, SLC9B2* and *VEGFA*) were computed in the results and more detailed information was displayed in the Supplementary Table S8. As annotated using OMIM databases, *MANBA* and *VEGFA* both fell into the entries directly related to kidney or diabetes mellitus. Furthermore, INTERPRO analysis was conducted for the protein annotation, these pleotropic genes were enriched into a variety of proteins, such as glycoside hydrolase family2 (*MANBA*), melatonin receptor family (*MTNR1B*), amino acid transporter, transmembrane (*SLC36A4*), glucose/ ribitol dehydrogenase (*SLC9B2*) and platelet-derived growth factor (*VEGFA*), which may be closely associated with eGFRcrea or T2D. Gene *VEGFA* is involved in the pathway of renal cell carcinoma, which is closely related to glucose transport and renal development generated by KEGG. All these results above furnish supporting evidence for our results from the functional aspect.

Discussion

In this study, we applied the cFDR approach to detect more pleotropic effects and tried to discover more of the missing heritability of two related traits/diseases. Using the pleotropic-informed statistical approach, we can improve the statistic power by 15–20 times to detect the non-null effects and genetic risk factors of both phenotypes compared to the

unconditional FDR (Andreassen et al. 2013c). Moreover, the analysis has clear advantage for correlated traits/diseases, as the method allows for the detection of risk loci regardless of their effect directions when compared to the traditional meta-analysis (Pei et al. 2014). Furthermore, this method enables identification of shared loci between eGFRcrea and T2D by leveraging the pleiotropic polygenic effects. To the best of our knowledge, we are the first to apply this method to detect common variants associated with both eGFRcrea and T2D.

By combining the largely independent datasets of eGFRcrea and T2D, we newly detected a total of 74 loci for eGFRcrea and 3 for T2D with the threshold of cFDR < 0.05, and we also identified four pleiotropic loci (rs227375 located in *MANBA*, rs1876602 located near *MTNR1B* and *SLC36A4*, rs4699049 located in *SLC9B2*, and rs9381257 located near *VEGFA* and *LINC01512*) associated with both eGFRcrea and T2D. The results demonstrated the feasibility of applying the pleiotropic-informed cFDR statistical approach to two related phenotypes: eGFRcrea and T2D. The novel loci detected by pleotropic effects enable us to have a better understanding of these two diseases and may provide some guidance for the subsequent studies of eGFRcrea and/or T2D.

The SNP rs4699049 mapping to solute carrier family 9 memberB2 (*SLC9B2*) is a newly identified association with both T2D and eGFRcrea in our study. Sodium/hydrogen exchanger *NHA2* is known as *NHEDC2* or *SLC9B2*, which is present in rodent and human β-cells (Deisl et al. 2013). The expression of *NHA2* plays a critical role in insulin secretion capacity of islets (Deisl et al. 2013). In *NHA2*-deficient mice, the defection of insulin secretion contributes to impaired glucose tolerance (Deisl et al. 2013). The *SLC9* gene family encodes Na^+/H^+ exchangers (NHEs), which are fatal for transepithelial movement of Na^+ and HCO^{3-} in the kidney (Fuster and Alexander 2014). Another study found that the distribution and expression of *NHA2* are especially limited to individual organs, such as in the bone or distal tubules of the kidney (Deisl et al. 2016). *NHA2* resides in the plasma membrane of Madin–Darby canine kidney (MDCK) cells, which are highly enriched and functionally significant in renal tubules (Chintapalli et al. 2015). To sum up, gene *SLC9B2* had been reported to be associated with T2D and kidney functions, but there was no earlier study which illustrated its relationship with both two traits.

Another pleiotropic SNP rs9381257 lay in the intergenic region near gene *VEGFA*, which encodes vascular endothelial growth factor A. On the one hand, gene *VEGFA* had been reported to have correlation with kidney in many aspects (Karihaloo et al. 2005; Eremina et al. 2007; Zeggini et al. 2008; Sharma et al. 2011). Gene *VEGFA* produced by renal podocytes is a requisite of glomerulogenesis and glomerular filtration barrier formation in animals (Eremina et al. 2007), and gene *VEGFA* can affect the number of nephrons by impacting ureteric bud growth during embryogenesis (Karihaloo et al. 2005). On the other hand, the earlier study demonstrated that this may be a signal for *VEGFA* associated with T2D ($p = 5 \times 10^{-6}$) in the discovery stage but displayed inconsistent evidence in the replication stage (Zeggini et al. 2008). Another published study indicated roles in the transcriptional modulation of *VEGFA* gene in T2D pathogenesis, and the signals near *VEGFA* probably increases susceptibility of T2D (Sharma et al. 2011). The functional annotation analysis for gene *VEGFA* revealed that it is associated with kidney development (GO 0001822) for the biological process and microvascular complications of diabetes 1

(OMIM: #603933). Despite extensive evidence demonstrating the relationship between gene *VEGFA* with eGFRcrea or T2D, no publication illustrated the association of *VEGFA* with both eGFRcrea and T2D.

The SNP rs1876602 was identified associated with both traits, which was located in the intergenic region near the genes *MTNR1B* and *SLC36A4*. A previous study has shown that the increase expression of *MTNR1B* is likely to decrease the release of insulin and lead to increasing risk of T2D (Tuomi et al. 2016). Another research has confirmed amphotericin B with copper (II) ions (AmB–Cu^{2+}) as a complex formed by *AmB* and is likely to cause upregulation of *MTNR1B* (Gola et al. 2015). In this study, it is illustrated that *AmB* is related to the oxidative injuries of the kidneys but studies about the influence of AmB–Cu^{2+} on renal cells are rare (Gola et al. 2015). Our evidence does not suggest a straightforward relationship between *MTNR1B* and eGFRcrea and T2D, but may provide a new insight for studying the common mechanism of eGFRcrea and T2D in genetic aspect.

In summary, our study showed pleiotropy effect between eGFRcrea and T2D by leveraging these two GWAS datasets with the cFDR analysis. First, we discovered several novel loci and confirmed some loci that had been identified in previous researches related to eGFRcrea or T2D. In addition, our results provided novel insights into shared genetic influences of eGFRcrea and T2D. Furthermore, we successfully expand the effective sample size for potential pleiotropic loci by detecting non-null effects and hence increasing the probability of these associations being replicated in the independent studies. We adopted cFDR to identify some novel loci that might explain more missing heritability of eGFRcrea or T2D, and effectively found four polygenic effects loci related to both eGFRcrea and T2D. However, our study may also have some limitations. First, it is unclear whether the correlation effect of two phenotypes in the joint consideration would shift the power of cFDR although this question might be partially addressed in future two-sample summary-based Mendelian Randomization (SMR) (Smith and Ebrahim 2004) study. Second, we could not provide information about the same ancestry analysis on the phenotypes (eGFRcrea and T2D) due to a lack of detailed individual-study-level data and it may have an impact of the results. Third, alternative approaches may be applied to check whether novel loci could still be identified to further confirm novel findings in our study or to furnish an empirical comparison of the relative performance of alternative methods, a topic we wish to pursue in a different area with comprehensive theoretical and simulation approaches. Most of our research results are significant at the statistical level as discovery, for which further clinical replications and further biological experiments are needed to further support our findings.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We appreciate the support from Zhengzhou University in providing necessary support for this collaborative project. HWD was partially supported by Grants from the National Institutes of Health [P50AR055081, R01AR057049, R01AR059781, D43TW009107, P20GM109036, R01MH107354, R01MH104680, R01GM109068 and R01AR069055], the Edward G. Schlieder Endowment fund to Tulane University.

References

Almgren P, Lehtovirta M, Isomaa B, Sarelin L, Taskinen MR, Lyssenko V, Tuomi T, Groop L. Heritability and familiality of type 2 diabetes and related quantitative traits in the botnia study. Diabetologia. 2011; 54:2811–2819. [PubMed: 21826484]

Andreassen OA, Djurovic S, Thompson WK, Schork AJ, Kendler KS, O'Donovan MC, Rujescu D, Werge T, van de Bunt M, Morris AP, et al. Improved detection of common variants associated with schizophrenia by leveraging pleiotropy with cardiovasculardisease risk factors. Am J Hum Genet. 2013a; 92:197–209. [PubMed: 23375658]

Andreassen OA, Thompson WK, Schork AJ, Ripke S, Mattingsdal M, Kelsoe JR, Kendler KS, O'Donovan MC, Rujescu D, Werge T, et al. Improved detection of common variants associated with schizophrenia and bipolar disorder using pleiotropy-informed conditional false discovery rate. PLoS Genet. 2013b; 9:e1003455. [PubMed: 23637625]

Andreassen OA, Thompson WK, Schork AJ, Ripke S, Mattingsdal M, Kelsoe JR, Kendler KS, O'Donovan MC, Rujescu D, Werge T, et al. Improved detection of common variants associated with schizophrenia and bipolar disorder using pleiotropy-informed conditional false discovery rate. Plos Genet. 2013c; 9:e1003455. [PubMed: 23637625]

Andreassen OA, McEvoy LK, Thompson WK, Wang Y, Reppe S, Schork AJ, Zuber V, Barrett-Connor E, Gautvik K, Aukrust P, et al. Identifying common genetic variants in blood pressure due to polygenic pleiotropy with associated phenotypes. Hypertension. 2014; 63:819–826. [PubMed: 24396023]

Chen J, Muntner P, Hamm LL, Jones DW, Batuman V, Fonseca V, Whelton PK, He J. The metabolic syndrome and chronic kidney disease in U.S. adults. Ann Intern Med. 2004; 140:167–174. [PubMed: 14757614]

Chintapalli VR, Kato A, Henderson L, Hirata T, Woods DJ, Overend G, Davies SA, Romero MF, Dow JAT. Transport proteins NHA1 and NHA2 are essential for survival, but have distinct transport modalities. Proc Natl Acad Sci USA. 2015; 112:11720–11725. [PubMed: 26324901]

Cho YS, Chen CH, Hu C, Long JR, Ong RTH, Sim XL, Takeuchi F, Wu Y, Go MJ, Yamauchi T, et al. Meta-analysis of genome-wide association studies identifies eight new loci for type 2 diabetes in east Asians. Nat Genet. 2012; 44:U67–U97.

Chung DJ, Yang C, Li C, Gelernter J, Zhao HY. GPA: a statistical approach to prioritizing GWAS results by integrating pleiotropy and annotation. Plos Genet. 2014; 10:e1004787. [PubMed: 25393678]

Cichonska A, Rousu J, Marttinen P, Kangas AJ, Soininen P, Lehtimaki T, Raitakari O, Jarvelin MR, Salomaa V, Ala-Korpela M, et al. metaCCA: summary statistics-based multivariate meta-analysis of genome-wide association studies using canonical correlation analysis. Genet Epidemiol. 2015; 39:540–540.

Cichonska A, Rousu J, Marttinen P, Kangas AJ, Soininen P, Lehtimaki T, Raitakari OT, Jarvelin MR, Salomaa V, Ala-Korpela M, et al. metaCCA: summary statistics-based multivariate meta-analysis of genome-wide association studies using canonical correlation analysis. Bioinformatics. 2016; 32:1981–1989. [PubMed: 27153689]

Cook JP, Morris AP. Multi-ethnic genome-wide association study identifies novel locus for type 2 diabetes susceptibility. Eur J Hum Genet. 2016; 24:1175–1180. [PubMed: 27189021]

Deisl C, Simonin A, Anderegg M, Albano G, Kovacs G, Ackermann D, Moch H, Dolci W, Thorens B, Hediger MA, et al. Sodium/hydrogen exchanger NHA2 is critical for insulin secretion in beta-cells. Proc Natl Acad Sci USA. 2013; 110:10004–10009. [PubMed: 23720317]

Deisl C, Anderegg M, Albano G, Luscher BP, Cerny D, Soria R, Bouillet E, Rimoldi S, Scherrer U, Fuster DG. Loss of sodium/hydrogen exchanger NHA2 exacerbates obesity- and aging-induced glucose intolerance in mice. Plos One. 2016; 11:e0163568. [PubMed: 27685945]

Efron B. Size, power and false discovery rates. Ann Stat. 2007; 35:1351–1377.

Eid A, Bodin S, Ferrier B, Delage H, Boghossian M, Martin M, Baverel G, Conjard A. Intrinsic gluconeogenesis is enhanced in renal proximal tubules of Zucker diabetic fatty rats. J Am Soc Nephrol. 2006; 17:398–405. [PubMed: 16396963]

Eremina V, Baelde HJ, Quaggin SE. Role of the VEGF-A signaling pathway in the glomerulus: evidence for crosstalk between components of the glomerular filtration barrier. Nephron Physiol. 2007; 106:32–37.

Fuster DG, Alexander RT. Traditional and emerging roles for the SLC9 Na+/H+ exchangers. Pflugers Arch Eur J Physiol. 2014; 466:61–76. [PubMed: 24337822]

Gola J, Skubis A, Sikora B, Kruszniewska-Rajs C, Adamska J, Mazurek U, Strzalka-Mrozik B, Czernel G, Gagos M. Expression profiles of genes related to melatonin and oxidative stress in human renal proximal tubule cells treated with antibiotic amphotericin B and its modified forms. Turk J Biol. 2015; 39:856–864.

Golay A, Ybarra J. Link between obesity and type 2 diabetes. Best Pract Res Clin Endocrinol Metab. 2005; 19:649–663. [PubMed: 16311223]

Greenbaum J, Wu K, Zhang L, Shen H, Zhang J, Deng HW. Increased detection of genetic loci associated with risk predictors of osteoporotic fracture using a pleiotropic cFDR method. Bone. 2017; 99:62–68. [PubMed: 28373146]

Huang DW, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res. 2009a; 37:1–13. [PubMed: 19033363]

Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc. 2009b; 4:44–57. [PubMed: 19131956]

Karaderi T, Drong AW, Lindgren CM. Insights into the genetic susceptibility to type 2 diabetes from genome-wide association studies of obesity-related traits. Curr Diab Rep. 2015; 15:83. [PubMed: 26363598]

Karihaloo A, Karumanchi SA, Cantley WL, Venkatesha S, Cantley LG, Kale S. Vascular endothelial growth factor induces branching morphogenesis/tubulogenesis in renal epithelial cells in a neuropilin-dependent fashion. Mol Cell Biol. 2005; 25:7441–7448. [PubMed: 16107693]

Kooner JS, Saleheen D, Sim X, Sehmi J, Zhang W, Frossard P, Been LF, Chia KS, Dimas AS, Hassanali N, et al. Genome-wide association study in individuals of South Asian ancestry identifies six new type 2 diabetes susceptibility loci. Nat Genet. 2011; 43:984–989. [PubMed: 21874001]

Kottgen A, Pattaro C, Boger CA, Fuchsberger C, Olden M, Glazer NL, Parsa A, Gao XY, Yang Q, Smith AV, et al. New loci associated with kidney function and chronic kidney disease. Nat Genet. 2010; 42:376–U334. [PubMed: 20383146]

LeBlanc M, Zuber V, Andreassen BK, Witoelar A, Zeng LY, Bettella F, Wang YP, McEvoy LK, Thompson WK, Schork AJ, et al. Identifying novel gene variants in coronary artery disease and shared genes with several cardiovascular risk factors. Circ Res. 2016; 115:83–94.

Li HX, Gan W, Lu L, Dong X, Han XY, Hu C, Yang Z, Sun L, Bao W, Li PT, et al. A genome-wide association study identifies GRK5 and RASGRP1 as type 2 diabetes loci in Chinese Hans. Diabetes. 2013; 62:291–298. [PubMed: 22961080]

Lorenzo C, Nath SD, Hanley AJG, Abboud HE, Gelfond JAL, Haffner SM. Risk of type 2 diabetes among individuals with high and low glomerular filtration rates. Diabetologia. 2009; 52:1290–1297. [PubMed: 19367385]

Lv WQ, Zhang X, Zhang Q, He JY, Liu HM, Xia X, Fan K, Zhao Q, Shi XZ, Zhang WD, et al. Novel common variants associated with body mass index and coronary artery disease detected using a pleiotropic cFDR method. J Mol Cell Cardiol. 2017; 112:1–7. [PubMed: 28843344]

Mahajan A, Go MJ, Zhang WH, Below JE, Gaulton KJ, Ferreira T, Horikoshi M, Johnson AD, Ng MCY, Prokopenko I, et al. Genome-wide trans-ancestry meta-analysis provides insight into the genetic architecture of type 2 diabetes susceptibility. Nat Genet. 2014; 46:234-+. [PubMed: 24509480]

Mahajan A, Rodan AR, Le TH, Gaulton KJ, Haessler J, Stilp AM, Kamatani Y, Zhu G, Sofer T, Puri S, et al. Trans-ethnic fine mapping highlights kidney-function genes linked to salt sensitivity. Am J Hum Genet. 2016; 99:636–646. [PubMed: 27588450]

Morris AP, Voight BF, Teslovich TM, Ferreira T, Segre AV, Steinthorsdottir V, Strawbridge RJ, Khan H, Grallert H, Mahajan A, et al. Large-scale association analysis provides insights into the genetic

architecture and pathophysiology of type 2 diabetes. Nat Genet. 2012; 44:981. [PubMed: 22885922]

O'Seaghdha CM, Fox CS. Genome-wide association studies of chronic kidney disease: what have we learned? Nat Rev Nephrol. 2012; 8:89–99.

Okada Y, Sim X, Go MJ, Wu JY, Gu D, Takeuchi F, Takahashi A, Maeda S, Tsunoda T, Chen P, et al. Meta-analysis identifies multiple loci associated with kidney function-related traits in east Asian populations. Nat Genet. 2012; 44:904–909. [PubMed: 22797727]

Parra EJ, Below JE, Krithika S, Valladares A, Barta JL, Cox NJ, Hanis CL, Wacher N, Garcia-Mena J, Hu P, et al. Genome-wide association study of type 2 diabetes in a sample from Mexico City and a meta-analysis of a Mexican-American sample from Starr County. Texas Diabetologia. 2011; 54:2038–2046. [PubMed: 21573907]

Pattaro C, Teumer A, Gorski M, Chu AY, Li M, Mijatovic V, Garnaas M, Tin A, Sorice R, Li Y, et al. Genetic associations at 53 loci highlight cell types and biological pathways relevant for kidney function. Nat Commun. 2016; 7

Pei YF, Zhang L, Papasian CJ, Wang YP, Deng HW. On individual genome-wide association studies and their meta-analysis. Hum Genet. 2014; 133:265–279. [PubMed: 24114349]

Peng C, Shen J, Lin X, Su KJ, Greenbaum J, Zhu W, Lou HL, Liu F, Zeng CP, Deng WF, et al. Genetic sharing with coronary artery disease identifies potential novel loci for bone mineral density. Bone. 2017; 103:70–77. [PubMed: 28651948]

Poulsen P, Kyvik KO, Vaag A, Beck-Nielsen H. Heritability of type II (non-insulin-dependent) diabetes mellitus and abnormal glucose tolerance-a population-based twin study. Diabetologia. 1999; 42:139–145. [PubMed: 10064092]

Regele F, Jelencsics K, Shiffman D, Pare G, McQueen MJ, Mann JF, Oberbauer R. Genome-wide studies to identify risk factors for kidney disease with a focus on patients with diabetes. Nephrol Dial Transpl. 2015; 30(Suppl 4):iv26–iv34.

Schork AJ, Thompson WK, Pham P, Torkamani A, Roddey JC, Sullivan PF, Kelsoe JR, O'Donovan MC, Furberg H, Schork NJ, et al. All SNPs are not created equal: genome-wide association studies reveal a consistent pattern of enrichment among functionally annotated SNPs. Plos Genetics. 2013; 9

Sharma NK, Langberg KA, Mondal AK, Elbein SC, Das SK. Type 2 diabetes (T2D) associated polymorphisms regulate expression of adjacent transcripts in transformed lymphocytes, adipose, and muscle from Caucasian and African–American subjects. J Clin Endocrinol Metab. 2011; 96:E394–403. [PubMed: 21084393]

Sim X, Ong RTH, Suo C, Tay WT, Liu JJ, Ng DPK, Boehnke M, Chia KS, Wong TY, Seielstad M, et al. Transferability of Type 2 diabetes implicated loci in multi-ethnic cohorts from Southeast Asia. Plos Genet. 2011; 7:e1001363. [PubMed: 21490949]

Sivakumaran S, Agakov F, Theodoratou E, Prendergast JG, Zgaga L, Manolio T, Rudan I, McKeigue P, Wilson JF, Campbell H. Abundant pleiotropy in human complex diseases and traits. Am J Hum Genet. 2011; 89:607–618. [PubMed: 22077970]

Smith GD, Ebrahim S. Mendelian randomization: prospects, potentials, and limitations. Int J Epidemiol. 2004; 33:30–42. [PubMed: 15075143]

Stahl EA, Wegmann D, Trynka G, Gutierrez-Achury J, Do R, Voight BF, Kraft P, Chen R, Kallberg HJ, Kurreeman FA, et al. Bayesian inference analyses of the polygenic architecture of rheumatoid arthritis. Nat Genet. 2012; 44:483–489. [PubMed: 22446960]

Tin A, Colantuoni E, Boerwinkle E, Kottgen A, Franceschini N, Astor BC, Coresh J, Kao WHL. Using multiple measures for quantitative trait association analyses: application to estimated glomerular filtration rate. J Hum Genet. 2013; 58:461–466. [PubMed: 23535967]

Tuomi T, Nagorny CLF, Singh P, Bennet H, Yu Q, Alenkvist I, Isomaa B, Ostman B, Soderstrom J, Pesonen AK, et al. Increased melatonin signaling is a risk factor for type 2 diabetes. Cell Metab. 2016; 23:1067–1077. [PubMed: 27185156]

Voight BF, Scott LJ, Steinthorsdottir V, Morris AP, Dina C, Welch RP, Zeggini E, Huth C, Aulchenko YS, Thorleifsson G, et al. Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis (vol 42, pg 579, 2010). Nat Genet. 2011; 43:388–388.

Yamamoto T, Nakagawa T, Suzuki H, Ohashi N, Fukasawa H, Fujigaki Y, Kato A, Nakamura Y, Suzuki F, Hishida A. Urinary angiotensinogen as a marker of intrarenal angiotensin II activity associated with deterioration of renal function in patients with chronic kidney disease. J Am Soc Nephrol. 2007; 18:1558–1565. [PubMed: 17409316]

Zeggini E, Scott LJ, Saxena R, Voight BF, Marchini JL, Hu T, de Bakker PIW, Abecasis GR, Almgren P, Andersen G, et al. Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. Nat Genet. 2008; 40:638–645. [PubMed: 18372903]

Zeng CP, Chen YC, Lin X, Greenbaum J, Chen YP, Peng C, Wang XF, Zhou R, Deng WM, Shen J, et al. Increased identification of novel variants in type 2 diabetes, birth weight and their pleiotropic loci. J Diabetes. 2016

Zhou R, Lin X, Li DY, Wang XF, Greenbaum J, Chen YC, Zeng CP, Lu JM, Ao ZX, Peng LP, et al. Identification of novel genetic loci for osteoporosis and/or rheumatoid arthritis using cFDR approach. PLoS One. 2017; 12:e0183842. [PubMed: 28854271]

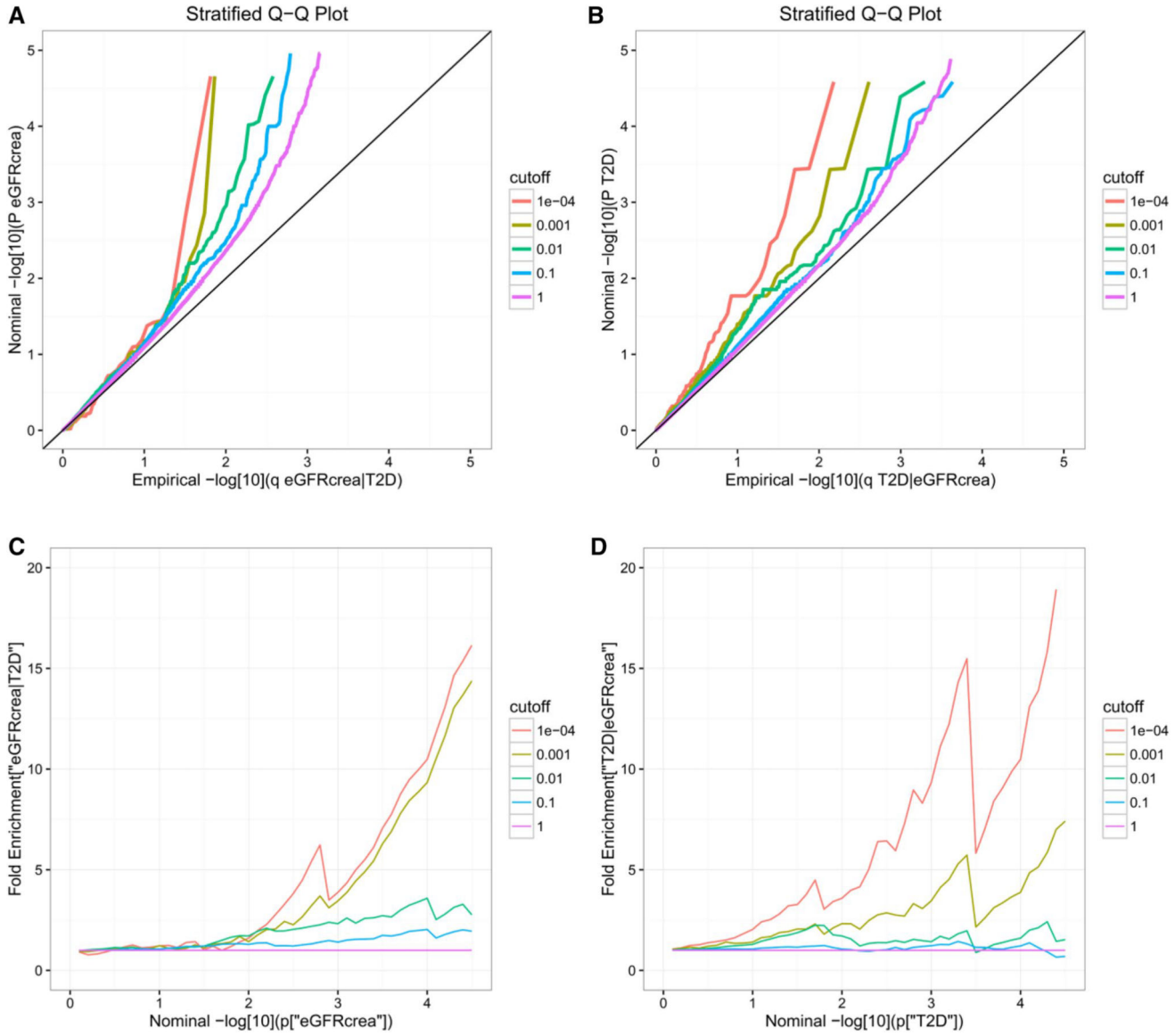

Fig. 1.

Stratified Q–Q (upper panel) and enrichment (lower panel) plots. Upper panel, stratified Q–Q plots of nominal versus empirical − $\log_{10}$ *p* values in (**a**) eGFRcrea as a function of significance of the association with T2D, and in (**b**) T2D as a function of significance of the association with eGFRcrea. Lower Panel: fold-enrichment plots of enrichment versus nominal − $\log_{10}$ *p* values for (**c**) eGFRcrea as a function of significance of the association with T2D, and (**d**) T2D as a function of significance of the association with eGFRcrea. The purple line with slope of zero represents all SNPs

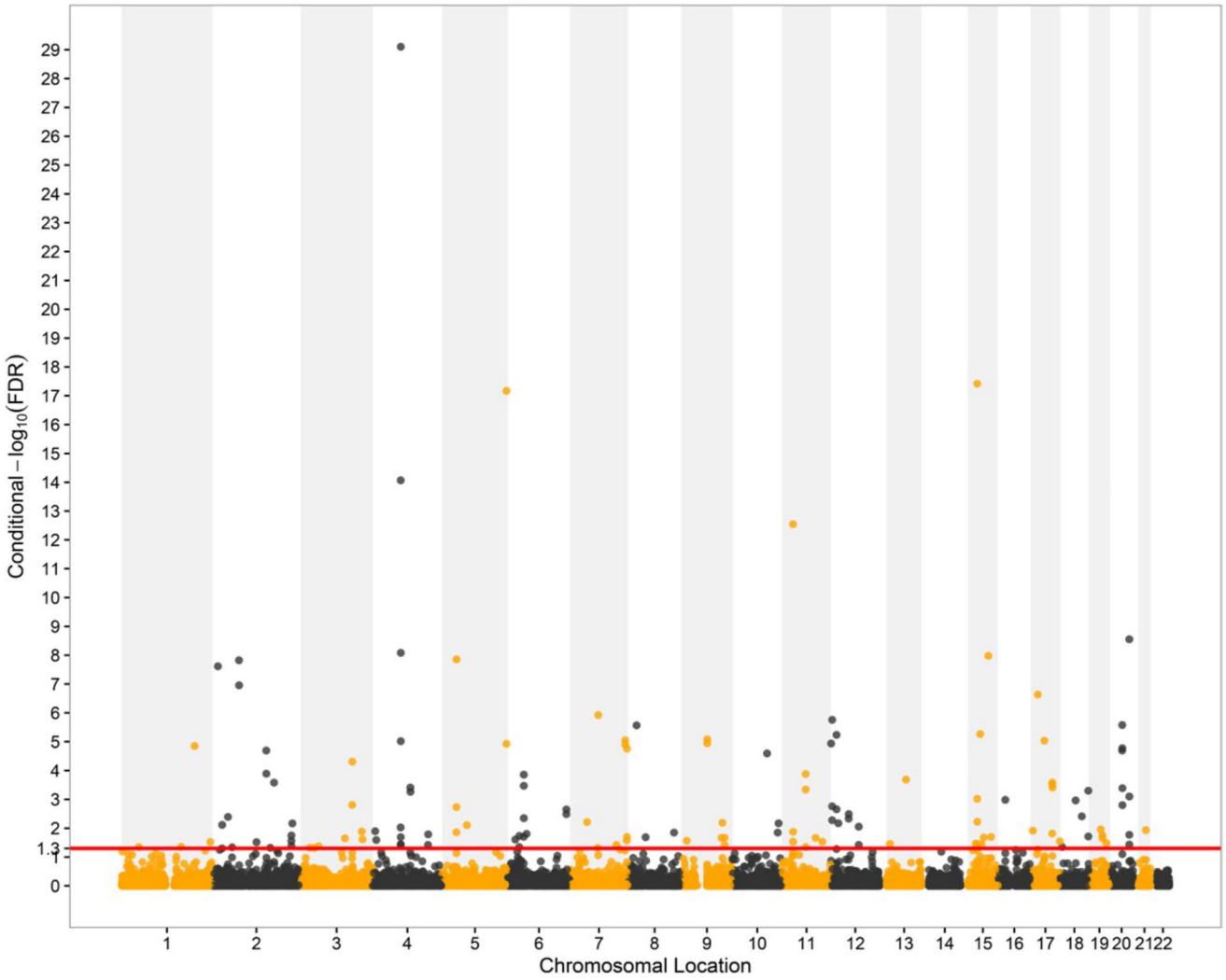

Fig. 2.

Conditional Manhattan plot of conditional − $\log_{10}$ FDR values for eGFRcrea given T2D (eGFRcrea |T2D). The red line marks the conditional − $\log_{10}$ FDR value of 1.3 which corresponds to a cFDR < 0.05

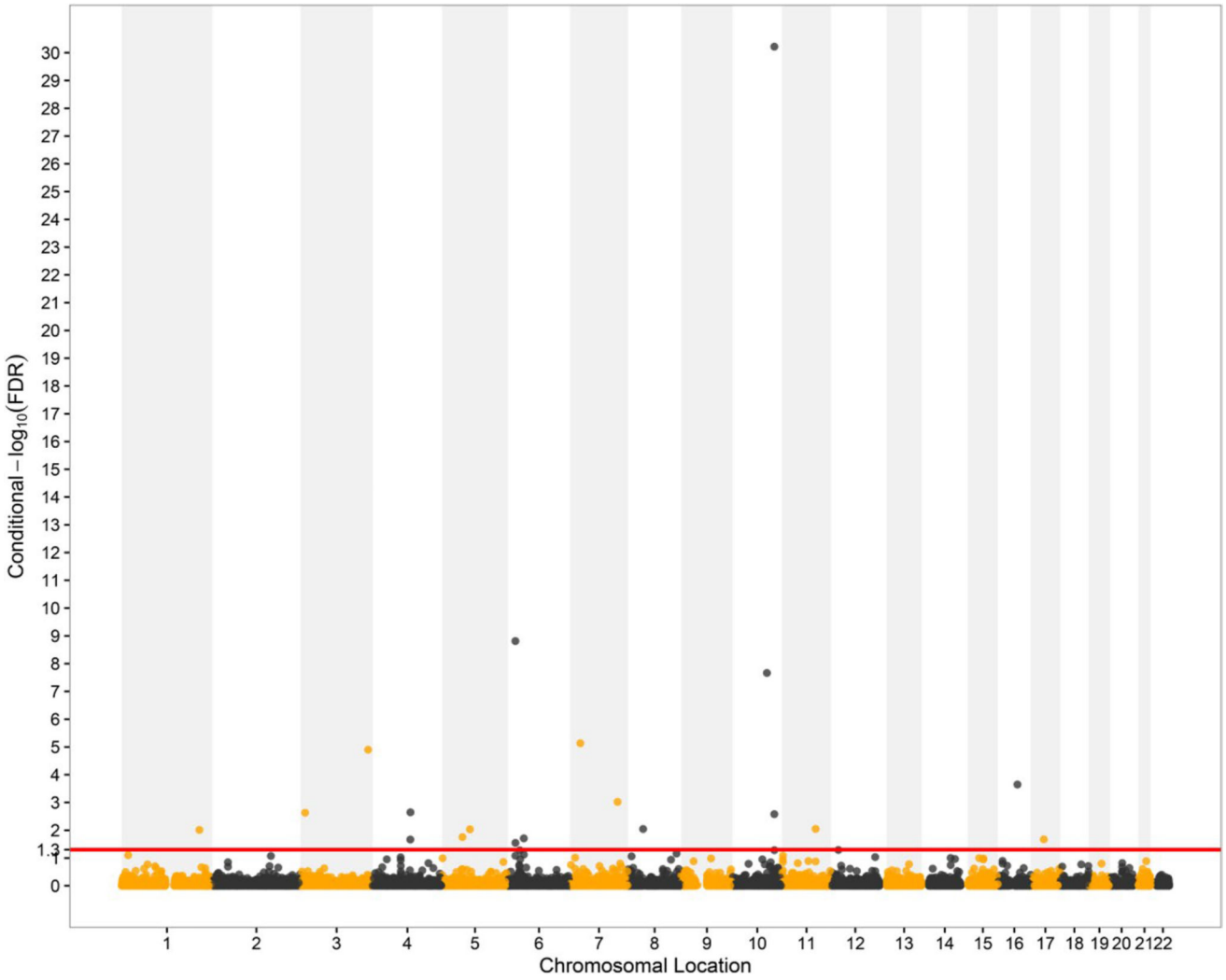

Fig. 3.

Conditional Manhattan plot of conditional − $\log_{10}$ FDR values for T2D given eGFRcrea (T2D| eGFRcrea). The red line marks the conditional − $\log_{10}$ FDR value of 1.3 which corresponds to a cFDR < 0.05

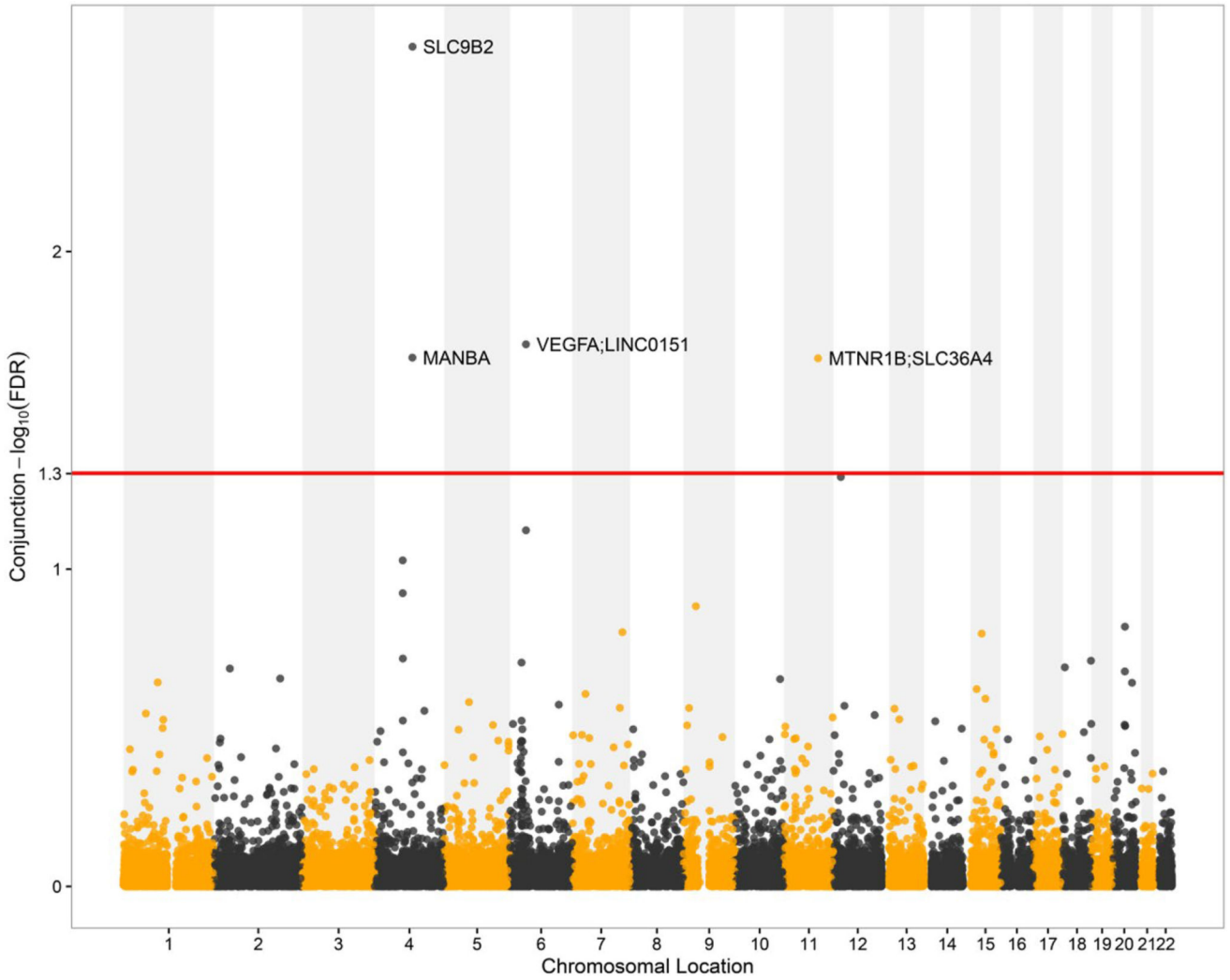

Fig. 4.

Conjunction Manhattan plot of conjunction − $\log_{10}$ FDR values for eGFRcrea and T2D. The red line marking the conditional − $\log_{10}$ FDR value of 1.3 corresponds to a ccFDR < 0.05

Table 1

Conjunction cFDR: pleiotropic Loci in eGFRcrea and T2D (cFDR < 0.05)

Locus	SNP	Role	Chr	Neighbor gene	cFDR		ccFDR
					eGFRcrea|T2D	T2D|eGFRcrea	
1	rs4699049	Intronic	chr4	SLC9B2[c]	0.00055	0.002262	0.002262
2	rs9381257	Intergenic	chr6	VEGFA[a] LINC01512[c]	0.000139	0.01961	0.01961
3	rs227375	Intronic	chr4	MANBA[b]	0.000389	0.0216	0.0216
4	rs1876602	Intergenic	chr11	MTNR1B[b] SLC36A4[c]	0.0217	0.008979	0.0217

[a]This gene is previously reported to be associated with eGFRcrea and T2D but the SNP is not

[b]This gene is previously reported to be associated with eGFRcrea or T2D but the SNP is not

[c]This SNP and gene have neither been reported to be associated with eGFRcrea and T2D